Rock weathering controls the potential for soil carbon storage at a continental scale

Supplementary Materials

Authors

Eric W. Slessarev[1]* (0000-0002-4076-1950), Oliver A. Chadwick[2] (0000-0002-4964-6203), Noah W. Sokol[1] (0000-0003-0239-1976), Erin E. Nuccio[1] (0000-0003-0189-183X), Jennifer Pett-Ridge[1]* (0000-0002-4439-2398)

Affiliations

[1] Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory

[2] Department of Geography, University of California Santa Barbara

*Co-corresponding authors: slessarev1@llnl.gov; pettridge2@llnl.gov

This file includes:

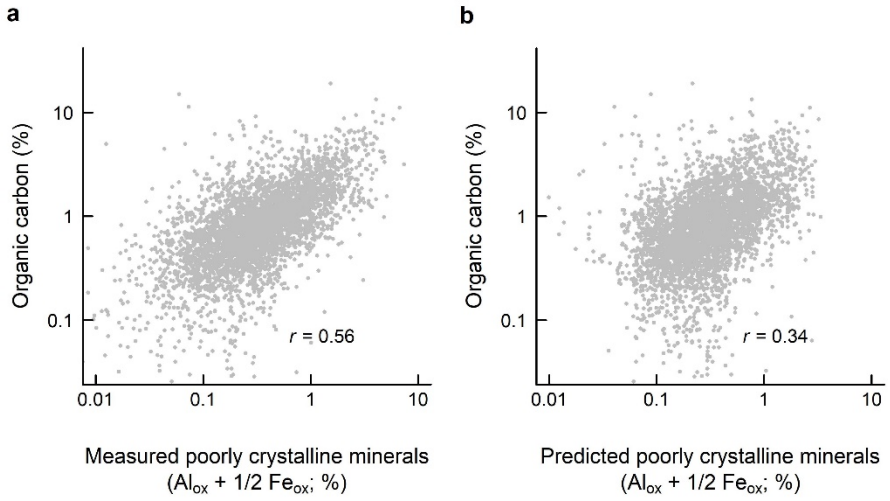
Supplementary Text 1: Soil profile calculation methods.

We obtained organic carbon data from the NCSS and RaCA databases. The RaCA database differs from the NCSS database in that it was compiled from a single sampling campaign using standard methods that covered the conterminous US evenly, and was intended primarily for quantification of SOC stocks – hence ancillary soil chemical variables (e.g. NH_4-oxalate extractable metals) are not reported. When possible, we used direct combustion analysis measurements to assign organic carbon data. When inorganic carbon was quantified, we calculated total organic carbon as the difference between total carbon and inorganic carbon. In the case of the NCSS database, inorganic carbon was not necessarily reported for all samples – in this case we used soil pH to infer the presence of inorganic carbon. When soil pH was less than 6.5 we assumed that no inorganic carbon was present. When soil pH was greater than 6.5 or unquantified and no inorganic carbon data were reported, we assumed that the samples potentially contained inorganic carbon. In this case, we used reported organic carbon values measured by wet oxidation when available. If no wet oxidation data were available, then organic carbon was assigned a missing value and the profile was not used. In the case of the RaCA dataset, samples were screened for inorganic carbon prior to quantification, and so we assumed that no inorganic carbon was present when it was not reported.

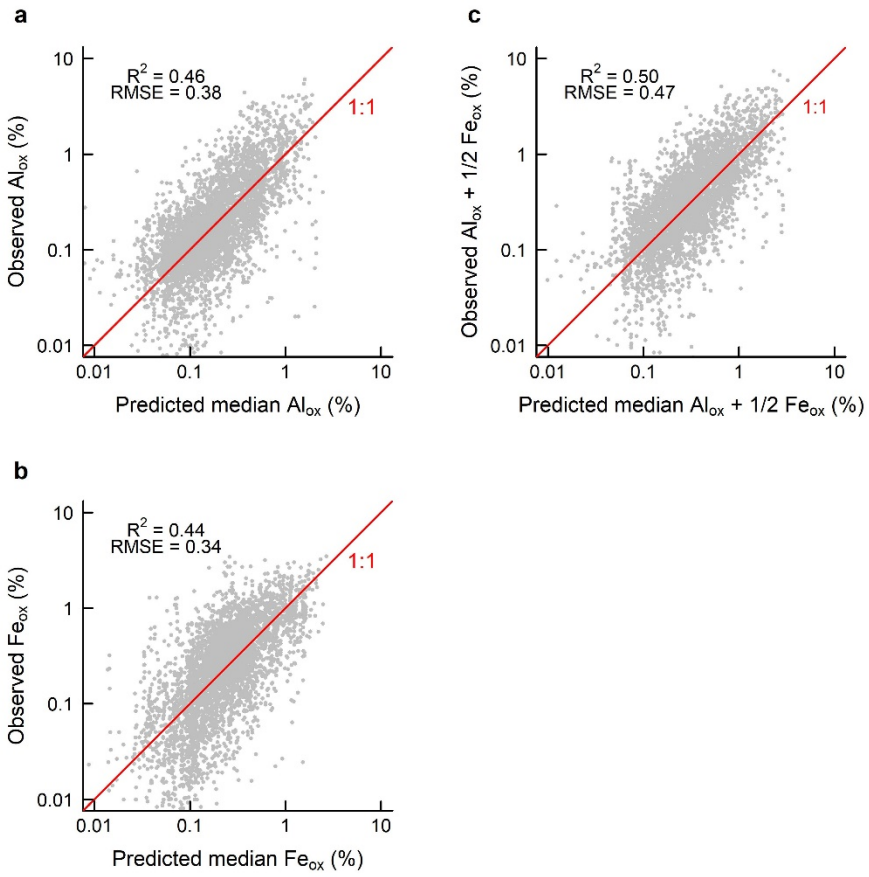
We summarized soil profile data by computing depth-weighted averages of % NH_4-oxalate extractable Al and Fe and % SOC for individual soil profiles. We did not compute stocks on an area-basis using bulk density measurements because this would have reduced the amount of data available (fewer profiles have measured density values). We excluded organic horizons when averaging profiles because the processes that control carbon accumulation in organic horizons are not strongly tied to mineralogy. Similarly, when analyzing data from RaCA, we

excluded soil profiles that were sampled from wetlands under the RaCA land-use classification scheme because the processes that control carbon storage in wetland soils are distinct. Organic horizons were identified using horizon nomenclature: all soil layers designated with a capital "O" or with SOC concentrations greater than 20 % were excluded. The remaining part of the profile was summarized by depth-weighted average to 100 cm, or to the depth of the lowermost C, AC, or BC horizon if the profile terminated above a depth of 100 cm. Profiles that terminated above 100 cm and did not include C, AC, or BC horizons were excluded, in addition to any profiles that were not contiguous.

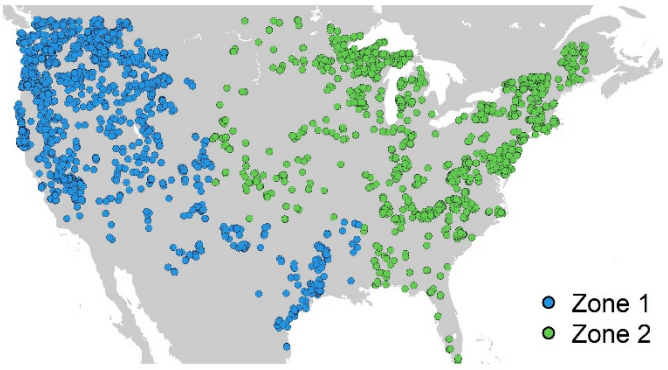
In a small number of cases NH_4-oxalate extractions were reported as exact zero values (n = 456 horizons for Al and 993 horizons for Fe; 1-2 % of the data). Similarly, when inorganic carbon was subtracted from total carbon to calculate organic carbon, organic carbon estimates were sometimes slightly negative (n = 644 NCSS horizons, 2% of the data). We adjusted these near-zero values by setting NH_4-oxalate extractable Al and Fe or organic C equal to one half the apparent lower limit of detection (0.01%) when reported values fell below this limit.



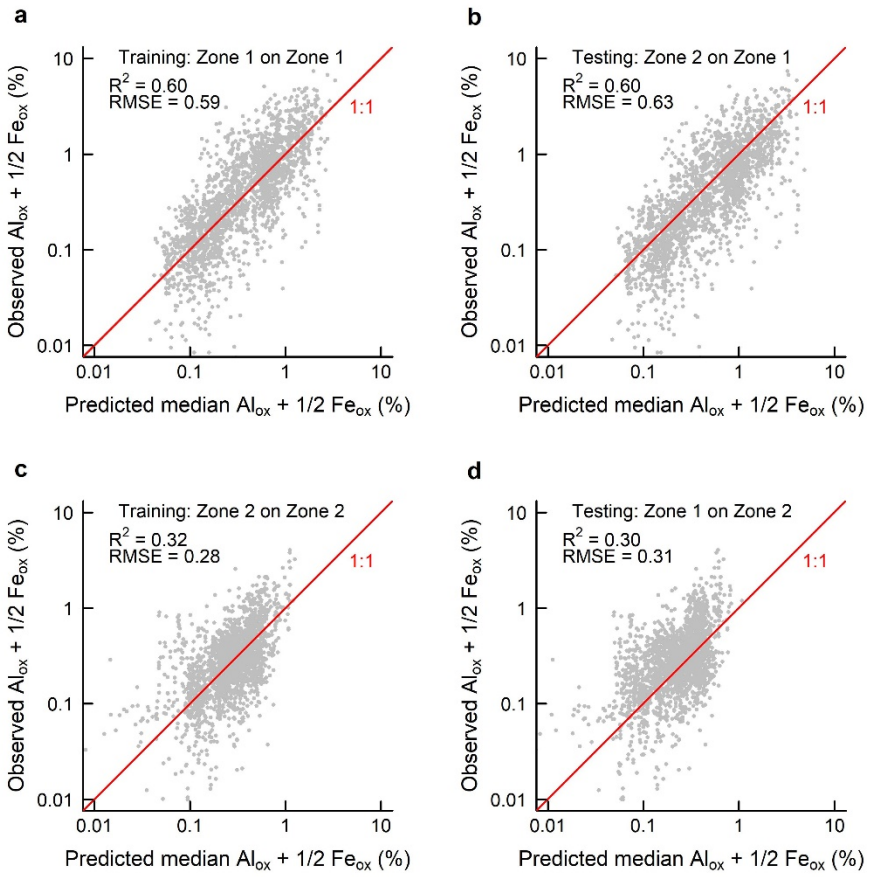
Supplementary Figure 1. Poorly crystalline minerals versus SOC across NCSS profiles. Panel (**a**) shows observed PCMs versus observed SOC across the 4,019 NCSS profiles with both PCM and SOC data included in this analysis, with each point representing the depth-weighted average of a single profile. Panel (**b**) shows the same data, but with predicted PCM abundance obtained from the statistical model shown on the horizontal axis. Pearson's correlation (*r*) between SOC and PCMs was 0.56 (95% CI [0.48, 0.62]) for measured PCMs and 0.34 (95 % CI [0.26, 0.41]) for predicted PCMs.



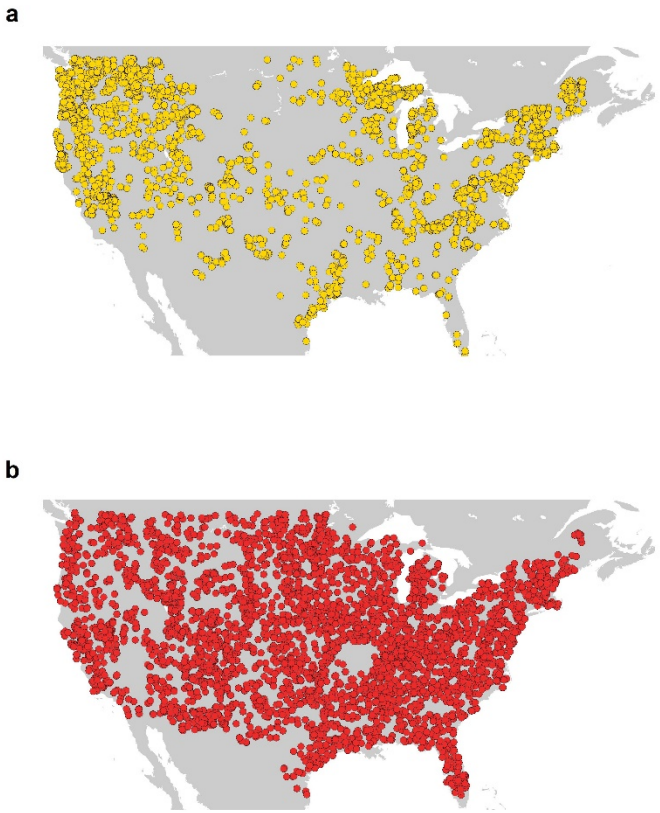
Supplementary Figure 2. Predicted PCMs versus observed PCMs. The three panels show data from the 4,019 profiles from the NCSS database that were used to train statistical models of PCM abundance, with predicted median depth-weighted averages on the horizontal axes and observed depth-weighted averages on the vertical axes. Panel (**a**) shows predicted Al_{ox} versus observed Al_{ox}, panel (**b**) show predicted Fe_{ox} versus observed Fe_{ox}, and panel (**c**) shows the weighted sum, Al_{ox} + ½ Fe_{ox}. Red lines show the 1:1 relationship, and R^2 values obtained by regressing log-transformed observations on log-transformed predictions are shown on each panel. Root mean squared error (RMSE) values were computed on un-transformed variables.



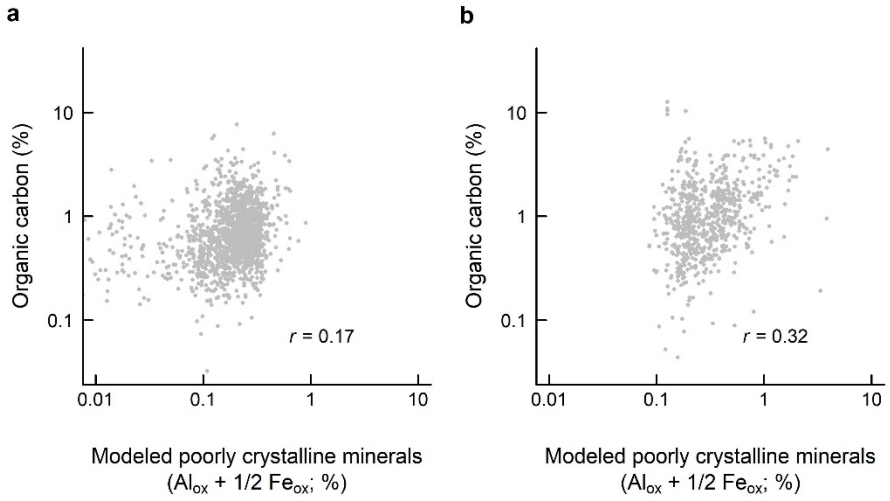
Supplementary Figure 3. Cross validation domains. Points are soil pits from the NCSS database used for model testing and training; Zone 1 is shown in blue, Zone 2 in green.



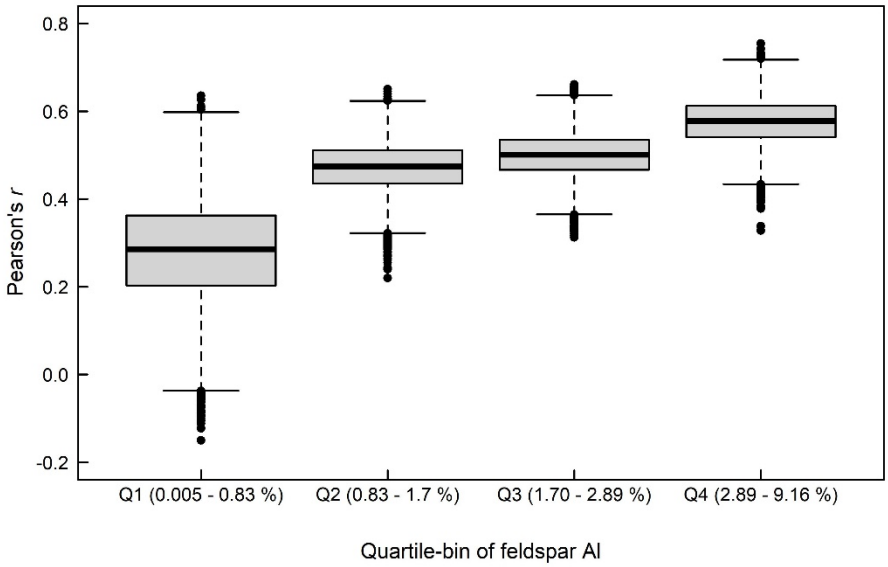
Supplementary Figure 4. Cross validation results. Panels (a) – (d) show predicted median Al_{ox} + ½ Fe_{ox} values versus observed Al_{ox} + ½ Fe_{ox}. Points represent depth-weighted averages of individual soil profiles from cross validation Zone 1 (panels **a, b**) and Zone 2 (panels **c, d**). Panel (**a**) shows predicted values for Zone 1 derived from a model trained on Zone 1, whereas (**b**) shows predictions derived from a model trained on Zone 2 and tested on Zone 1. Panel (**c**) includes predicted values for Zone 2 from a model trained on Zone 2, and (**d**) predictions from a model trained on Zone 1 and tested on Zone 2. Red lines show the 1:1 relationship, and R^2 values obtained by regressing log-transformed observations on log-transformed predictions are shown on each panel. Root mean squared error (RMSE) values were computed on un-transformed variables.



Supplementary Figure 5. Geographic distribution of soil profiles included in this analysis. Panel (**a**) shows soil profiles from the NCSS database with oxalate extraction and SOC data (yellow points). Panel (**b**) shows soil profiles from the Rapid Carbon Assessment (RaCA) database that were included in this analysis (red points). Profiles were excluded if the total thickness of mineral horizons was < 100 cm and no data from the C horizon were reported.



Supplementary Figure 6. Correlation between predicted PCMs and SOC across RaCA. Points represent depth-weighted average of individual profiles reported in the RaCA dataset, which gives relatively even coverage across the conterminous United States. Depth-weighted average poorly crystalline mineral abundance (Al_{ox} + ½ Fe_{ox}) was inferred using statistical models (see Methods). (**a**) RaCA data from the weathered domain (MAP ≥ PET and primary Al ≤ 1.7 %); (**b**) RaCA data from enhanced weathering zones (MAP ≥ PET and primary Al > 1.7 %). Pearson's correlation coefficient (*r*) was 0.17 in the weathered domain (95% CI [0.09, 0.24]), and 0.32 inside enhanced weathering zones (95% CI [0.20, 0.42]).



Supplementary Figure 7. Correlation between PCMs and SOC versus primary Al. Box plots show the distribution of estimates of Pearson's correlation coefficient across subsets of the data defined by binned values of feldspar Al (on the horizontal axis), only including data where MAP ≥ PET and MAP - PET < 1 m. Correlations were computed using spatial blocks to weight and sample the data (see Methods). Distributions represented by the box plots were obtained by resampling blocks with replacement 10,000 times and recalculating the correlation coefficient. Center lines of the boxes show medians; box ends show the 1st and 3rd quartiles; whiskers show extremes.

Coefficient	Zone 1 value [95 % CI]	Zone 2 value [95 % CI]
$Al_{ox} = a_1*Q*Al_{pri} + a_2*Q*Al_{sec} + a_3*Al_{tot}$		
a_1	$8.5*10^{-5}$ [$2.1*10^{-5}$, $12.4*10^{-5}$]	$13.5*10^{-5}$ [$10.7*10^{-5}$, $16.7*10^{-5}$]
a_2	$4.4*10^{-5}$ [$2.0*10^{-5}$, $9.5*10^{-5}$]	$2.2*10^{-5}$ [$1.5*10^{-5}$, $3.0*10^{-5}$]
a_3	0.011 [0.009, 0.013]	0.012 [0.008, 0.015]
$Fe_{ox} = b_1*Q*Fe_{pri} + b_2*Q*Fe_{sec} + b_3*Fe_{tot}$		
b_1	$11.4*10^{-5}$ [$5.3*10^{-5}$, $19.6*10^{-5}$]	$32.6*10^{-5}$ [$22.0*10^{-5}$, $40.6*10^{-5}$]
b_2	$7.0*10^{-5}$ [$2.8*10^{-5}$, $12.1*10^{-5}$]	$5.4*10^{-5}$ [$1.6*10^{-5}$, $8.9*10^{-5}$]
b_3	0.034 [0.028, 0.043]	0.048 [0.039, 0.070]

Supplementary Table 1. Quantile regression statistics from cross validation. Values are fitted coefficients for each model, fitting the 50^{th} quantile. The 95% confidence intervals were obtained from spatially blocked bootstrapping (see Methods). The variable Q was in units of mm when fitting the model. Values for cross validation zone 1 are in the left column, and values for cross validation zone 2 are in the right column.